

62914

IDRC - Lib

62914

ISM85 - 33

Information
Sciences
Archival Copy

EVALUATION OF BIOTECHNOLOGY DATA BASES

71414

Bev Chataway

May 1985

Added
drawn
no. 01

doc 447

In response to a request from Martha Stone to review the data bases currently available in the area of biotechnology, the following report was submitted.

Trying to define biotechnology is itself a problem. It is a truly multi-disciplinary field covering a wide range of disciplines (biology, chemistry, agriculture, energy, pharmaceuticals, water and industry, food and feedstuffs). Thus its literature comes in a variety of formats and is scattered across all subjects. The popularity of this concept and the resulting explosion of articles, reports and patents in this area has exacerbated the problem of literature searching of biotechnology.

Data bases specializing in biotechnology have only been in existence for a short time. Just two years ago, only DERWENT BIOTECHNOLOGY ABSTRACTS and TELEGENLINE were available as international public access systems. Today, reference sources and vendors' catalogues identify 4 other data bases focusing on this topic: ABSTRACTS IN BIOCOMMERCE (ABC), CURRENT BIOTECHNOLOGY ABSTRACTS, BIOTECHNOLOGY INVESTMENT OPPORTUNITIES and BIOBUSINESS.

The following paragraphs outline the features for each data base, their strengths and weaknesses. Appendix I reproduces sample records from 4 of these specialized tools.

DERWENT BIOTECHNOLOGY ABSTRACTS - is available online only to subscribers at the cost of about \$500 per year. This subscription fee covers the cost of the printed monthly abstract and permits access to the data base which costs about \$100/hour to search. Over 9,000 entries are included in this file with about 600 new records added each month. Entries date from January 1983. All aspects of biotechnology are covered. Approximately 1000 of the world's leading scientific and technical journals as well as conference reports and patents are covered. It is available through Orbit. The main drawbacks of this data base are: lack of coverage of commercial news reports, the subscription fee and its small size.

TELEGENLINE - Based on the printed publication Telegen Reporter, this data base covers scientific journal and report literature and commercial news material. It is reported that this is the only data base to cover some biotechnology conferences without published proceedings. Some patents (US only) are included from April 1983. Records do go back to 1973 but only in small numbers. Abstracts are included for entries since March 1982. Abstracts include a statement of the document type (research report, news feature etc). Descriptive phrases are added. A strength of this data base is its ability to provide full texts of some documents. It is supposed to be particularly strong in genetic engineering. Updated monthly, there are more than 6000 entries in the file. No subscription fee is charged however on-line access costs \$95 per contact hour. Available on Dialog and ESA. The Head of Reference at CISTI believes that this data base is one of the best available in this subject area.

ABSTRACTS IN BIOCOMMERCE - is currently available only in the United Kingdom and concentrates on the trade press especially newsletters. Few abstracts of patents or scientific research are included. Abstracts are fairly short. Entries begin with March 1981 and number about 4000. Descriptors are not added. This data base concentrates mainly on genetic engineering and monoclonal antibodies. Private file service is available. Shortcomings of this data base are its small size, limited accessibility and scarceness of scientific research.

CURRENT BIOTECHNOLOGY ABSTRACTS - Begun in May 1983 by the United Kingdom's Royal Society of Chemistry, this file is available through the European Space Agency. It covers about 200 scientific and trade sources. The abstracts are split principally into TECHNOCOMMERCIAL DATA (ie business news), TECHNIQUES (genetic manipulations, fermentation etc), INDUSTRIAL AREAS (pharmaceuticals, agriculture, energy) and INFORMATION (books, reviews). The file size numbers about 7000 references with monthly updates around 400 references. Short abstracts are prepared directly from the original sources but additional abstracts if available are taken from entries in other data bases eg. BIOLOGICAL ABSTRACTS. Controlled terms and a unique classification system are other features. With an hourly rate of only \$75 this data base is reasonably priced. Its only weakness is its lack of a large backfile for retrospective searches.

BIOTECHNOLOGY INVESTMENT OPPORTUNITIES - is a full text data base containing the full text of Biotechnology Investment Opportunities, a newsletter on investment opportunities in genetic engineering. Company profiles, current research and government policies are included. Coverage is International but with little developing country focus. It is available through Newsnet.

BIOBUSINESS - On April 15, 1985, Biosciences Information Service (BIOSIS) and the Information Access Company (IAC) announced the creation of BIOBUSINESS. It is designed to provide information to business executives and product development people on the business opportunities of biotechnology research. Over 1100 serials will be scanned. The data base will contain over 28,000 records. Available this summer on major data base vendors, this file will provide fully indexed references and abstracts.

To compare these data bases, two topics

- 1) cell and tissue culture as applied to agriculture
- 2) fermentation as applied to antibodies, vaccines and tropical diseases

were tested on DERWENT BIOTECHNOLOGY ABSTRACTS, TELEGENTLINE and CURRENT BIOTECHNOLOGY ABSTRACTS. The other data bases were not accessible either through IDRC's Library or CISTI. In addition the same searches were performed on a number of other data bases covering agriculture, health and science to ascertain overall coverage of these topics.

Search terms were selected from Mr. Hulse's review paper on biotechnology. The search strategies used appear in Appendix II.

TOPIC ICELL AND TISSUE CULTURE IN AGRICULTURE

Data Base	No. of Hits	Developing Country Coverage	Language	Comments
DERWENT BIOTECHNOLOGY ABSTRACTS	907	Some	All languages represented	Very technical studies only.
TELEGENLINE	288	Some	Some foreign language material included	Popular and technical materials covered. Short abstracts.
CURRENT BIOTECHNOLOGY ABSTRACTS	68	Very little	Much foreign language material included especially Japanese	Very short abstracts. Controlled subject terms are difficult to search.

OTHER USEFUL DATA BASES

TOPIC I CELL AND TISSUE CULTURE IN AGRICULTURE

Data Base	No. of Hits	Developing Country Coverage	Abstracts	Host	Language	Entries From	Comments
CAB	3989	Yes	Yes	DIALOG ESA CAN/OLE	Mainly English	1972	SDI Available
AGRICOLA 10 110	3237 308	Some but primarily American	No	BRS DIALOG ORBIT	Some foreign language materials	1970	SDI Available
SCI SEARCH 1984-85 1981-83 1978-80	48 88 77	Very few	No	DIALOG	Primarily English	1974	Current but expensive. SDI available.
TROPAG	214	Good coverage	Yes	ORBIT	70% English Materials	1975	Very useful-Long Abstracts.
NATIONAL NEWSPAPER INDEX	22	None	No	DIALOG	English only	1979	Popular American press.
AGRIS	2319	Yes	Some	ESA	All languages	1975	SDI Available
LIFE SCIENCES COLLECTION	1397	Some coverage	Yes	DIALOG	Lots of foreign materials included	1978	Biological emphasis

OTHER USEFUL DATA BASES cont'd

TOPIC I CELL AND TISSUE CULTURE IN AGRICULTURE

Data Base	No. of Hits	Developing Country Coverage	Abstracts	Host	Language	Entries From	Comments
PASCAL	835	Name	Yes	QUESTEL	French and English	1976	Covers a wide range of subjects including bio-technology
IALINE	43	Almost none	Short Abstracts	QUESTEL	English and French	1970	
FAIREC	426	Since it is on tropical and sub-tropical fruits and vegetables, majority of records deal with developing countries	Short Abstracts	QUESTEL	60% English 15% French	1970	Useful.
BIOSIS 1981-85 1977-80 1969-76	3118 1969 701	Some coverage	Yes extensive indexing	BRS OLE CAN/ DIALOG ESA	Some foreign language materials	1969	SDI available. Basic research
NTIS	262	Almost none	Yes	BRS CAN/OLE DIALOG ORRITT ESA	Primarily English	1964	Long abstracts. Primarily U.S. government reports
CONFERENCE PAPERS INDEX	590	A bit.	No	DIALOG ESA	English	1973	Covers over 1000 conferences annually.

TOPIC IIFERMENTATION AND VACCINES, ANTIBODIES

Data Base	No. of Hits	Developing Country Coverage	Language	Comments
DERWENT BIOTECHNOLOGY ABSTRACTS	57	Some	All languages represented	Very technical studies. Oriented to basic research only.
TELEGENLINE	124	Some but very little	Primarily English language materials	News releases as well as technical reports are included. short abstracts.
CURRENT BIOTECHNOLOGY ABSTRACTS	8	Very little	Rate of foreign language materials	Very short abstracts. Patents are included. Controlled terms include inversions which are difficult to search.

TOPIC II

FERMENTATION AND VACCINES, ANTIBODIES ETC

Data Base	No. of Hits	Developing Country Coverage	Abstracts	Host	Language	Entries From	Comments
Medline							
1983-85	24	Some	Yes	BRS	Some	1966-	Inexpensive.
1980-82	22			DIALOG	German,		Extensive indexing
1977-79	13			NLM	Polish,		system
1975-76	18			BLAISE	Russian		
1971-74	40				items		
1966-70	9				served		
BIOSIS PREVIEWS							
1981-85	100			BRS	English	1969-	Biological
1977-80	38	Few items related to developing countries	Yes extensive indexing	CAN/OLE			emphasis
1969-76	20			DIALOG			
EMBASE (Excerpta Medica)							
1982-85	632			DIALOG	Lots of foreign language materials	1975-	Medical emphasis
1980-81	26	Very little	Yes				
1975-79	69						
NTIS	32	Almost none	Yes	BRS	Primarily English	1964	Long abstracts. Primarily U.S. government reports.
				CAN/OLE			
				DIALOG			
				ESA			
				ORBIT			
LIFE SCIENCES COLLECTION	118	Very little	Yes	DIALOG	Lots of foreign materials included.	1978	Biological emphasis especially GENETICS.

TOPIC II

FERMENTATION AND VACCINES, ANTIBODIES ETC cont'd

Data Base	No. of Hits	Developing Country Coverage	Abstracts	Host	Language	Entries From	Comments
SCISEARCH							
1984-85	5	None	No	DIALOG		1974-	Very current but fairly expensive. Document delivery service.
1981-83	2						
1978-80	0						
1974-77	0						
PASCAL	41	A bit.	Yes	QUESTEL ESA	French and English	1973	
INTER-NATIONAL PHAR-MACEUTICAL ABSTRACTS	3	None	Yes	BRS NLM DIALOG	Yes	1970	
CA SEARCH							
1982-85	84	Very little	Yes	BRS CAN/OLE DIALOG ORBIT QUESTEL	Lots of foreign language material included	1967	Very technical.
1980-81	9						
1977-79	9						
1972-76	18						
1967-71	8						
NATIONAL NEWSPAPER INDEX	15	None	No	DIALOG	English only	1975	Popular American press
CONFERENCE PAPERS INDEX	2	A bit	No	DIALOG ESA	English	1973	Covers over 1000 conferences annually.

Although the new specialized data bases promise comprehensive coverage of biotechnology, as the results of the searches demonstrate, they are by no means comprehensive in scope or coverage. The older data bases provided access to more references than the 3 specialized sources. With these specialized data bases there is overlap and duplication in some areas (scientific literature found in the core journals) and spotty coverage in others (developing country literature). No single data base can be used as the only source. In fact it has been recommended in data base searching literature that searchers working on biotechnology related questions should search several sources to ensure the most complete results eg. both TELEGENLINE and DERWENT BIOTECHNOLOGY ABSTRACTS. In addition, being so new, these specialized data bases do not have the extensive backfiles available from other sources like BIOLOGICAL ABSTRACTS (1969-), MEDLINE (1966-) and CAB (1973-).

Other factors to be considered when looking at these data bases included:

- 1) document delivery - how easily could researchers in developing countries obtain copies of listed documents
- 2) language of publication - in many of the data bases, much of the literature is in Japanese, Russian and German - do researchers in developing countries read these languages or have easy access to translation services
- 3) costs - access to these data bases is not cheap
- 4) variety of hosts - in searching the test questions, no single host supplied all relevant sources. The European Space Agency, Orbit and Dialog were all accessed.
- 5) the largest proportion of materials covered by these data bases focused on scientific research. Very few documents referred to ethical, safety or legal aspects of biotechnology
- 6) developing country coverage - in almost every case developing country material was scarce (exceptions CAB, FAIREC, and TROPAG).

My comments up to this part have focused on only bibliographic data bases however patent data bases, sequence databanks and culture collections are all important sources for biotechnology.

There are no patent data bases devoted solely to biotechnology however as previously mentioned DERWENT BIOTECHNOLOGY ABSTRACTS and CURRENT BIOTECHNOLOGY ABSTRACTS do include basic patents. WPI and the INPADOC files are also useful.

DNA and protein sequence data banks do exist. GENBANK and the European Molecular Biology Laboratory Nucleotide Sequence Data Library are two sources.

The area of culture collections of micro-organisms or cell lines will likely be an area of data base growth in the future. For instance CODATA has created a data base entirely on hybridomas ("a cell culture consisting of a clone of fused (hybrid) cells of different kinds") and monoclonal antibodies. Its name is CODATA-IUIS HYBRIDOMA DATA BANK: A DATA BANK ON CLONED CELL LINES AND THEIR IMMUNOREACTIVE PRODUCTS. On-line access is not yet available but requests can be submitted for searches. There is no charge for this service. Coverage is international. The data base has 151 searchable fields and contains approximately 3000 records.

New information sources for biotechnology are constantly being created. For example, just recently Eugene Garfield's Institute for Scientific Information produced the ISI - Atlas of Science: Biotechnology and Molecular Genetics. Britain's Science Reference Library is currently working on a one year project "European Biotechnology Information Project" to evaluate information needs and resources of the European Community in this area. This project involves an on-line system for placing enquiries with the SRL and may also include the creation of data bases concerned with the Library's patent and book holdings.

This short paper is not an in-depth analysis of biotechnology data bases but I hope it will provide some insight into the variety of sources available. I would be pleased to answer any questions or provide clarification or additional information or any of the topics covered.

SAMPLE RECORDS FROM THE BIOTECHNOLOGY DATA BASES

DERWENT BIOTECHNOLOGY ABSTRACTS (ON ORBIT)

-1-

- AN - 85-03985
CL - L (En) E4
TI - In vitro plantlet regeneration from juvenile and mature *Quercus robur* L.
ST - Plant propagation
AU - Vieitez A M; San-Jose M C; Vieitez E
LO - Plant Physiology, CSIC, Apartado 122, Santiago de Compostela, Spain.
JL - JHSCAS; J.Hortic.Sci.; (1985) 60, 1, 99-106
IW - English oak seedling and stump sprout propagation, culture medium effect
AB - Embryos and shoot apices from 3-4 month old seedlings and stump sprouts of *Quercus robur* were used to establish shoot cultures in medium containing 1 mg/l benzyladenine (BA). Shoot tips of initial cultures of mature trees were subcultured, and best shoot multiplication was observed using Heller's medium with 1 mM ammonium sulfate, or Greshof and Doy's medium. BA at 0.1 mg/l enhanced shoot growth. The basal ends of in vitro regenerated shoots were dipped in 0.5-1% w/l indolebutyric acid (IBA) for 2-10 min, and then transferred to auxin-free agar medium. Rooting rates of 83% and 63% were obtained from juvenile- and adult-derived shoots respectively. (29 ref)

ABSTRACTS IN BIOCOMMERCE

Celltech Ltd
Sumitomo Corporation

*40000688

Celltech has appointed Sumitomo as its exclusive agent in Japan. The company plans to make a significant impact in Japan on the diagnostics market over the short and medium term with monoclonal antibodies in kits using the IRMA system and other assays, which will be developed by the company in the future. The Japanese diagnostics market is estimated at £200M. Sumitomo is also selling Celltech's anti alpha interferon for purification and assays and, currently, Celltech is developing products to purify and assay beta and gamma interferons. Sumitomo is to work towards introducing Celltech's ABO blood grouping reagents in Japan and will promote Celltech's contract research services, which include recombinant DNA technology, gene synthesis, metabolic conversions, genome mapping, hybridoma technology and the contract production of monoclonal antibodies.

Celltech Press Release OCT82

Scrip 25OCT82

European Chemical News 01NOV82

Guardian 19OCT82

Daily Telegraph 20OCT82

Biotechnology News 15NOV82

Observer 14NOV82

BioEngineering News 01NOV82

Genetic Technology News DEC82

Genetic Engineering Letter 10DEC82

CURRENT BIOTECHNOLOGY ABSTRACTS (ON ESA)

0440020 7712 0937.420711.4111 350024=

1. *Chlorophyll a* (Chl *a*)

10-10-1964

Document type: JOURNAL Article Category: Electric-

Chem. Inc., Denver, CO 80202, 303-841-7561; FAX: 303-841-3068.

CODE: 1-14

05-91-3-1481-107

Class: object is method return

1. The following information is provided for the year ended 31 December 2014:

APPENDIX I

TELEGENLINE (ON DIALOG)

PHYTOGEN REGENERATES COTTON PLANTS FROM TISSUE CULTURE.
GENETIC ENGINEERING NEWS, BASE ON DIALOG, NO. 5, 1985.
PHYTOGEN RESEARCHERS HAVE REPORTED THE FIRST SUCCESSFUL REGENERATION OF
COTTON PLANTS FROM TISSUE CULTURE. THE ABILITY TO SUCCESSFULLY GROW WHOLE
PLANTS FROM UNDIFFERENTIATED CELLS COULD LEAD TO IMPROVED STRAINS, EITHER
THROUGH SOMOCLONAL VARIATION, OR THROUGH GENETIC ENGINEERING. THE COTTON
MARKET CURRENTLY IS WORTH \$4.5 BILLION. THE FIRM HAS APPLIED FOR A PATENT
IN

KEYWORD LISTING FOR TEST QUESTIONS

QUESTION I

Set 1

tissue culture
plant cell culture
cell tissue
clone*
somatic cell*
haploid*
biotechnology

Set 2

plant*
flower*
crop*

Combine sets 1 and 2.

Additional terms in French were used as PASCAL, IALINE and FAIREC.

QUESTION II

fermentation
biopharmaceutics
biotechnology

antibodies
vaccines
tropical medicine
tropical diseases
names of individual diseases

Combine sets 1 and 2.

Additional terms in French were used as PASCAL.